

EMPOWERING POTENTIALS OF OPEN DATA?
A STUDY EXAMINING INDIVIDUAL OPEN DATA
UTILIZATION IN FINLAND

Aleksi Ruuhilahti

University of Tampere
Faculty of Communication Sciences
Information Studies and Interactive Media
Master's Thesis
November 2018

In the recent years the amount of open data has proliferated over the internet. The term refers to data which is licensed in such way that everyone can access, reuse and redistribute it and by far the biggest distributors of open data have been governmental and municipal agencies around the world. As administration has started to share their data for the use of citizens and enterprises, this process is expected to produce versatile social and economic value for the society.

In this study I have focused on exploring the social potentials of open data and examined with qualitative approach how utilization of open data can empower the Finnish utilizers from individual and citizen's perspective. Based on the data that I collected with online-survey and four theme-interviews, it can be argued that open data utilization empowers the utilizers particularly by expanding their individual consciousness and awareness (power from within) and data utilization can be seen to strengthen utilizer's individual voice as citizen. The findings related to citizen influence or effective participation actualizing with the help of open data utilization however were limited among my results and thus the access to data alone do not seem to foster citizen's power to act in the society.

Keywords: Open data, open government data, empowerment, e-democracy, citizen participation

Contents

1	INTRODUCTION.....	1
2	CONCEPTUALIZING OPEN DATA AND DATA UTILIZATION	4
2.1	What is open data?.....	4
2.2	Studies examining open data utilization at raw-data level	6
2.2.1	Processes with data and popular outcomes	7
2.2.2	The common utilizer and motives to utilize	8
2.3	Open data and citizen participation in democratic govern	11
3	EMPOWERMENT THEORY AND OPEN DATA	14
3.1	Defining the concept.....	15
3.2	How different types of power help to understand empowerment?.....	16
3.3	A multidimensional process	18
3.4	Open data utilization, an unlike approach of empowerment research.....	20
4	STUDY AIMS AND METHODOLOGY	21
4.1	Research problem and the aims for the study.....	21
4.2	Research approach.....	22
4.2.1	Data collection methods	22
4.2.2	Thematic structure for interviews.....	24
4.2.3	Implementation of survey and interviews	25
4.2.4	Qualitative data analysis.....	26
5	ANALYSIS OF THE SURVEY AND INTERVIEWS	28
5.1	Respondents' affiliations with open data and perspectives on its potential ...	28
5.2	Themes of individual utilization.....	30
5.2.1	Data to understand	30
5.2.2	Optimizing with data	32
5.2.3	Data as utility for skill development and self-expression.....	33
5.2.4	Goal-oriented vs. experimental data use.....	34
5.2.5	The thematization in comparison with earlier research.....	36
5.3	Experiences regarding benefits and impacts of open data utilization	37
5.3.1	Social influence as the by-product of experimental utilization	38
5.4	Individual competence defines data utilization	41
5.4.1	Factors hampering open data utilization.....	43
6	DISCUSSION & CONCLUSIONS	45
6.1	Empowering potentials of open data utilization	45
6.2	The limitations of my study and suggestions for further research	48
	REFERENCES	50

APPENDIX 1: Online-survey in Finnish

1 INTRODUCTION

The coming of ICT, internet and digital culture has expanded the faith in stronger democracy. Also modern digital technologies, such as social media, Wikipedia and the likes at the head have always been accompanied with the promises of non-hierarchy, inclusive and equal participation possibilities and wider distribution of power and knowledge as they have been developed. Although these promises must be considered critically and not taken for granted, it can be regarded that new technologies are indeed increasing the access to information for wider group of people and sharing of information is easier than ever. With the help of information the quantity of knowledge which precedes understanding and wisdom is also assumed to grow among people and widespread information, knowledge and wisdom are thus intrinsic parts of ideal democratic societies.

In this light the recently proliferated amount of *open government data* (“data or content that is produced or commissioned by government and can be freely used, reused and re-distributed by anyone”, <http://opendefinition.org/government/>) seems interesting phenomena and, for one’s part, can be regarded as another promise for stronger democracy and citizen participation in the digital era. Like Rob Kitchin (2014, 9-10) notes in his book, data forms the base for the so called knowledge pyramid and has significant latent value that is realized when data is converted to information and knowledge. When the formerly exclusive access to administration data is opened for the use of citizens and enterprises, the novel uses are expected to realize versatile value for the whole society.

At the moment the known impacts and benefits regarding the use of open data are, however, in large extent only speculative and the empirical research concerning the topic is scarce. Being inspired by the promises given for open data (see Janssen et al., 2012) and particularly those related to social benefits, I am willing to shed light on this field of research by focusing on the individual utilization of open data in Finland. In their study of emerging impacts of open data, Davies et al. (2013) identified “inclusion and empowerment” as one of the themes where the use of open data is supposed to bring changes. “Participation and self-empowerment of citizens” is also one of the several benefits of open data identified by Janssen et al. (2012). By following these notions, in this thesis my aim is to examine how different individual uses of open data are related to these emphasized themes and analyze how utilization of open data empowers individual utilizers to

act on situations that affect their lives. The examination and discussion considering the question of empowerment is based on the following sub-questions:

- How open data is used from individual perspective and what are the purposes for utilization?
- What kind of benefits and impacts data utilization produces?
- What kind of factors enable data use among utilizers?

The development of different applications and services that are based on open data are emphasized in the discussions of open data but in my view it is important to notice that data can also offer multiple other utilization possibilities and empower individuals in contexts still scarcely known. Therefore the aim of this research is to explore open data utilization from wide perspective without limiting the scope of possible observations to concern only certain group of people or some already recognized utilization scenarios. Due to the wide perspective of my approach this study can be regarded as explorative and open for new observations concerning novel phenomenon. The analysis is based on individual experiences of open data utilizers and example cases of data utilization that I have collected with online-survey and 4 theme interviews. I'm approaching my research questions from qualitative perspective and I use qualitative content-analysis as my method to examine the collected data. The type of my analysis can be regarded as theory-bound as I aim to produce the units of analysis from the data itself but I'm still utilizing the earlier research concerning the phenomenon and theories of empowerment.

The structure of this thesis is divided in six different chapters. In chapter 2 my first aim is to conceptualize the meaning for the term *open data* so that the term can be used for the purposes of my research. In chapter 2 I also present dimensions of open data utilization in the light of earlier research and discuss how open data and data utilization can be seen to relate to concepts like citizen participation and e-democracy. In chapter 3 I present different definitions for the concept of empowerment and discuss how these theories of empowerment can help to understand what kind of potentials are embedded in open data. In chapter 4 I open up the research problem and my research questions more closely and discuss the research methods that I have chosen to answer my questions. In chapter 5 I present the results of my analysis which are based on the data gathered with the survey and interviews. The aim of chapter 5 is to create a foundation for the discussion of my main research question regarding the empowering potentials of open data utilization. In chapter 6 I discuss these potentials in the light of my results and present my evaluation

considering the success of my study. In chapter 6 I also present some possible directions for the further research.

2 CONCEPTUALIZING OPEN DATA AND DATA UTILIZATION

In this chapter I aim to conceptualize and discuss terms that are at the core of my study. In chapter 2.1 I conceptualize the meaning for the term open data so that the term can be used for the purposes of my study. In chapter 2.2 I present dimensions of open data utilization in the light of earlier research and in chapter 2.3 I discuss how open data utilization can be seen to relate to concepts like citizen participation and e-democracy.

2.1 What is open data?

Rob Kitchin (2014, 1) states that “data are commonly understood to be the raw material produced by abstracting the world into categories, measures and other representational forms that constitute the building blocks from which information and knowledge are created.”. Data is also perceived as the base for knowledge pyramid (Figure 1), a continuum which represents the common understanding considering formation of knowledge and eventually wisdom. In this study I regard the term *data* in the same commonly manner as Kitchin (2014, 1) and use it to refer to elements like numbers, characters or symbols which form the bedrock for the knowledge pyramid. The term data in this study, however, can also sometimes be mixed with the term information as there is ongoing debate how the concepts of data, information and knowledge are understood.

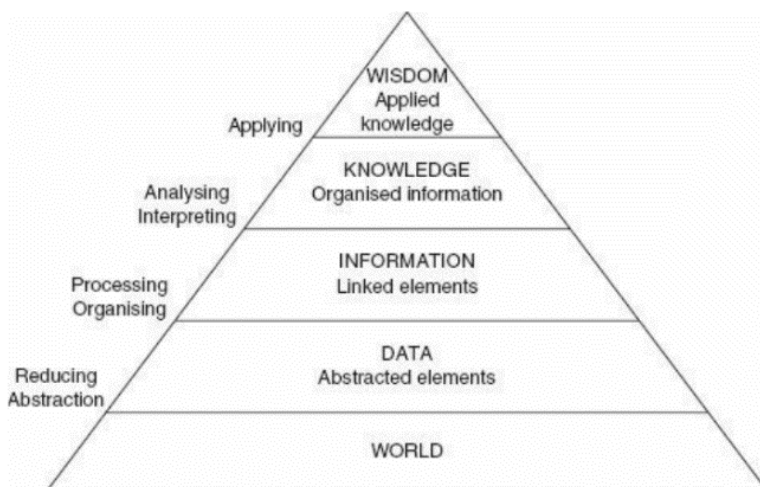


Figure 1. Knowledge pyramid (adapted from Kitchin 2014)

When discussing the concept of open data, the same debate between data, information and knowledge seems apparent. In the literature open data is often understood as raw, digital material which everyone can access, reuse and redistribute. In their webpage Open Knowledge Foundation sums up the definition for open data as following: “Open data and content can be freely used, modified, and shared by anyone for any purpose” (<http://opendefinition.org/>). Open data can, however, be seen to belong under the umbrella of wider concept, *open knowledge*. Open Knowledge Foundation continue their definition by claiming that open data are the building blocks of open knowledge and transformation from open data to open knowledge can be actualized when data is useful, usable and used (<https://okfn.org/opendata/>). The definition seems interesting from the standpoint of my research since I’m willing to study how open data is used and therefore examine the effects that the knowledge originating from the data can produce.

Open data can be produced and distributed by any actor but so far the biggest distributors have been governments, public administration agencies and municipalities around the world. In Finland the overall situation concerning the availability of open data and the offered possibilities for its use seems quite good when compared globally. In the recent years Finnish public sector has actively invested in the development of effective open data infrastructure and data opening as a general theme is highlighted in the current government program by Prime Minister Sipilä (<https://vm.fi/en/opendata>). With the help of these investments, major Finnish open data portals like avoindata.fi and HRI.fi have been launched and several municipalities over the country are running their own open data programs. The latest Global Open Data Index, an annual measurement of the publication of open government data maintained by Open Knowledge Network, also ranked Finland in the fifth place in the global comparison between countries (<https://index.okfn.org/place/>).

In the cases where the data and information that is collected, produced and disseminated by public sector bodies is usually referred with the term *Public Sector Information (PSI)*. When this data is distributed under an open license, it can be called *open government data* or *OGD* (<https://www.europeandataportal.eu/en/providing-data/goldbook/open-data-nut-shell>.) This study focuses particularly on the utilization and potentials contained in open government data although there can be references to data utilization that is not originated from governmental data sources. The focal point of my research is also on the utilization

and utilizers of “raw open data” in Kitchins’s (2014, 1) sense. In other words, I’m interested in studying how data, that is made open and accessible through data portals, Application programming interfaces (APIs) and the likes, is individually utilized. It must be noted that open data can also be utilized at end-user level for instance in cases where data is already processed in the next form and used indirectly through web-services, mobile applications or other compositions of data. However, this study is not focusing on the data utilization at end-user level.

2.2 Studies examining open data utilization at raw-data level

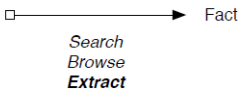
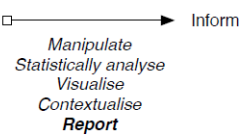
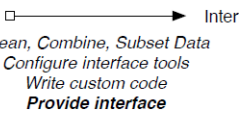
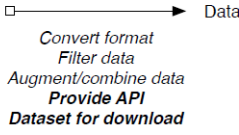
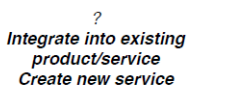
Due to the novelty of the phenomena, the amount of studies exploring the use of open data from citizens’ perspective is scarce. Several authors, like Gurstein (2011), Janssen et al. (2012) and Johnson (2014) criticize the myth around open data for being instantly empowering for everyone and emphasize “data divide” between those who actually can utilize the data and those who can’t. Despite of their apt critique against the optimistic promises for open data these works however do not pay attention to the question of how empowerment is experienced by those “empowered” to utilize data. Open data research in Finland is also quite tenuous except the works of Heli Koski and others (2015, 2017) which consider the impacts of open data from the economic and business perspective.

The earlier research examining open data use from social perspective seems to be focused on cases hailing from US and UK (Davies 2010, Kassen 2013, Okamoto 2016), countries which can be seen as forerunners of open data. These studies consider data utilization also from the raw-data perspective. The study of Davies’ (2010) is by far one of the most comprehensive case studies exploring the topic. By examining 55 instances of data use from data.gov.uk -portal and using mixed-method research design Davies aims to explore the possible implications of OGD use for different models of democratic change and public sector reform. Another relevant look at the use cases of open data is Karen Okamoto’s “What is being done with open government data? An exploratory analysis of public uses of New York City open data.” (2016). Aspiring to study how open data use “promotes innovative strategies for social progress” Okamoto gathered a total of 77 apps, visualizations, websites and blogs utilizing open data from City’s Open Data Portal as a basis for her analysis.

2.2.1 Processes with data and popular outcomes

According to Davies (2010) the outcomes of data use from UK's government open data portal (data.gov.uk) are diverse and Davies states that many of these appear to be exploratory or experimental in their nature. By using mixed data gathering and analysis methods, starting from participant-observation and ending to in-depth interviews and embedded cases, Davies identified five distinct processes of OGD use. The processes, explanations describing each processes' characteristics and the number of instances in Davies' data are presented in Table 1. Davies points that the processes can overlap and many uses employ multiple processes. This matter can explain the difference between the instances presented in Table 1 (74) compared to the 55 instances that acted as a base for Davies' analysis.

Table 1. Five processes of OGD use identified by Davies (2010)

Process (n=instances)	Summary (and example)
Data  Fact (n=8)	A dataset is used directly to identify a specific fact of interest. E.g. Finding out the voting history of a local constituency.
Data  Information (n=19)	Content from a dataset is given a single representation or interpretation that is reported in text or graphics. E.g. Composing a report that "profile[s] communities of interest within [the local area] as part of the Council's equality & diversity agenda".
Data  Interface (n=26)	An interface is provided allowing interactive representation of a dataset – providing information customized to the user's input. E.g. Creating a searchable interactive online map of stations and former British rail assets.
Data  Data (n=17)	A derivative dataset is provided for download, or access via an API E.g. I "took Westminster Constituency data, combined it with scraped [General Election] 2005 data and exposed it as RDF."
Data  Service (n=4)	A service is provided that relies on open data, whilst not necessarily exposing it to the end-user. E.g. Using boundary data from the Census to run an application that forwards reports of Potholes to the correct Highways authority.

Davies' (2010) findings indicate that the process of data-to-interface appears to be the most popular way to utilize data in the context of data.gov.uk. Also the processes that aim at creating static information from data, like graphs or texts composed based on data, seem popular.

The study of Okamoto's (2016) also offers views regarding data use. Instead of analyzing the data use as processes Okamoto has specified the artifacts created from data based on their type or format. From total of 77 items Okamoto identified 48% of them being in a form of some kind of map. The other formats or types that Okamoto found from her sample were web apps (19%), infographics (9%), data extractors (6%), mobile apps (5%), APIs (2%) and charts (2%).

When comparing the findings of Davies' (2010) and Okamoto's (2016) studies, the results of OGD use from two different countries seem to be quite similar. Okamoto's notion about maps being the most used format when building something from New York's open data associates with Davies' findings about the popularity of data-to-interface and data-to-information processes in the UK. Okamoto tells a little about the characteristics of the maps analyzed so practically map as a format can fit in to multiple processes described by Davies. Also the amount of web- and mobile-apps among Okamoto's results relates to Davies' notions about data-to-interface being the most used process of open data use.

2.2.2 The common utilizer and motives to utilize

Around the field of open data there seems to be an observable discourse of "data-for-developers", a fact that is emphasized in several writings and studies (see Lassinantti et al. 2014, Gurstein 2011). The results of empirical studies exploring the data utilizers (Davies 2010, Okamoto 2016) support the idea that skilled developers or professionals with technical or quantitative backgrounds are the ones interested to work with open data. Based on his study, Davies (2010) describes the common utilizer being a highly educated male working either in public sector or private SME firms. The informants in Davies' study associated themselves strongly with labels like 'IT Specialist', 'Data Specialist' and 'Web Developer' but also 'Citizen' (Davies, 2010). The observations from US made by Okamoto (2016) follow the same line with no radical anomalies. Her findings suggest that the utilizers are working either individually or in groups and "individuals with technical backgrounds" were the most popular type of utilizer (61%) among her sample. In addition to individual developers, Okamoto also identified government departments

(13%), tech companies (9%) and non-profit civic organizations (8%) as utilizers of open data.

The findings above demonstrate that technological understanding, high education and special skills associate strongly with the current use and utilizers of open data. Based on the data outcomes explored by Davies (2010) and Okamoto (2016), it is also obvious that the creators of the most popular use cases (data-to-interface or data-to-maps and web-applications) most likely need to know how to code or analyze data to some extent. Although “technicality” could be described as a ruling nominator around the use of open data Davies (2010) still emphasizes the diversity of utilization possibilities that open data offers. The amount of data-to-fact, data-to-information and data-to-data processes appearing in his results (Table 1) indicate that there exists use cases that differ from application development and cases where high technical understanding is not necessarily required in order to gain something from the data.

The focus on technology is also apparent as Davies (2010) explored data utilizers’ motivations to engage with open government data. Based on respondents’ self-assessed motivations, Davies identified six overlapping motivational clusters for OGD use: government focused, technology & innovation focused, reward focused, digitizing government focused, problem solving and social or public sector entrepreneurialism. The motivational clusters, their relationships and appearance on two-dimensional axis are presented in Figure 2.

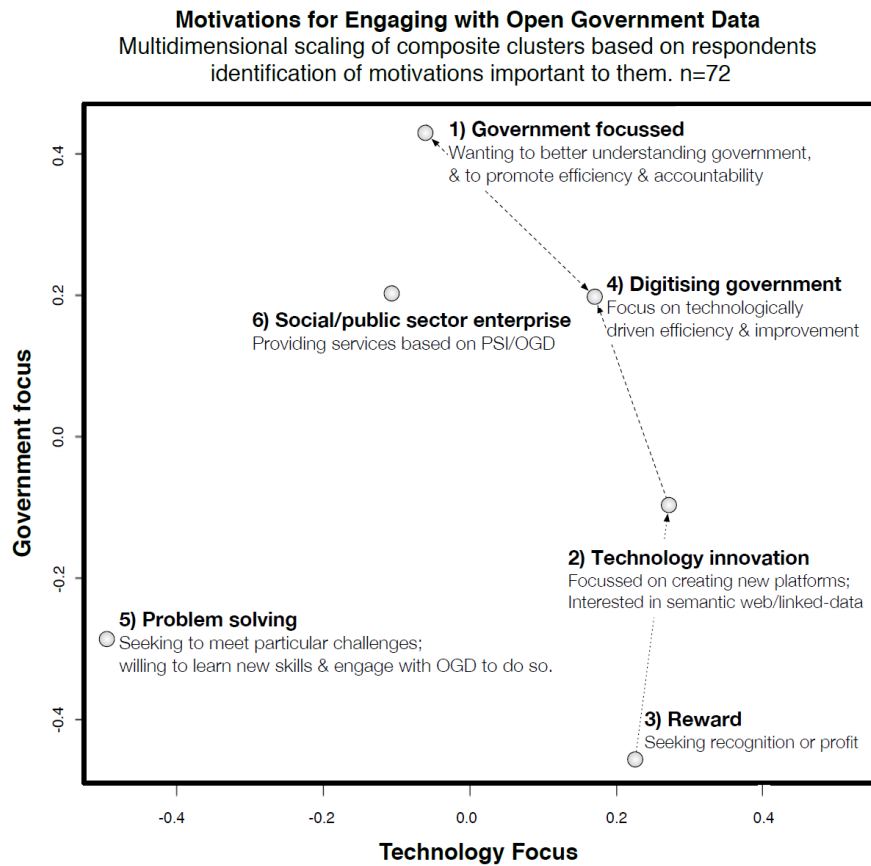


Figure 2. Motivational clusters of open government data use identified by Davies (2010)

The most significant cluster in Davies' (2010) results was the cluster labelled as "Digitizing government". According to Davies this cluster contains beliefs of technology driven reform of government's efficiency and functioning where utilization of OGD plays a major role. Davies also states that in the most cases individuals are driven by multiple goals and motivations and clusters 1, 2 and 3 are closely interrelated. The arrows in Figure 2 indicate the correlation between the clusters but also draw on the common relationships between the clusters based on qualitative data (Davies 2010).

Based on his analysis considering OGD use from data.gov.uk and the utilizers' motivations to engage with OGD, Davies (2010) identified total of 9 distinct types of civic OGD use (Table 2). According to Davies, these types indicate the different models how the use of OGD supports democratic engagement and public sector reform.

Table 2. Civic uses of OGD by Davies (2010)

(a) Political participation	(b) Collaborative / community based participation	(c) Individual choice / market participation
a1. Supporting scrutiny a2. Informing citizens as voters a3. Informing citizens on specific issues a4. Supporting campaigning and lobbying	b1. Co-produced planning b2. Co-produced services b3. Co-produced information	c1. Improving the quality of demand (better informed consumers) c2. Improving the quality of provision (competitive innovation)

2.3 Open data and citizen participation in democratic govern

Distribution of government data is one part of development towards more openly functioning governance, sometimes referred as *open government* (<http://www.oecd.org/gov/open-government.htm>). In a larger scale, I see that open government data is also linked with the discussion of *digital democracy* or *e-democracy*. E-democracy and the same sort of concepts are based on the idea of a new, more civic model of democracy which builds from the nature of information society. (Anttiroiko 2003, 27) Data and information that administration and different agencies are collecting and managing are crucial parts of their operation and different policies at municipal and state level are based on this data. As the access to this basic commodity is shared forward to citizen level, it is believed, at least rhetorically, to strengthen democracy in general and possibilities of citizens to participate and influence on societal issues.

In his article which considers citizen participation and influence within information society Anttiroiko (2003, 14) defines *participation* as opportunities for citizens or groups to be involved in processes in which important collective issues are being discussed or executed. According to Anttiroiko (2003, 14), participation can also be considered as more loose action in which person is involved in some sort of social process with other people. The concept of participation seems however not that unambiguous and Anttiroiko (2003, 16-17) regards that other contiguous concepts related to participation are needed in order to understand participation more closely in theoretical sense. According to Anttiroiko (2003, 16-17) these concepts are *social engagement* and *influence*. Social engagement describes individual's adherence to community and is strongly based on indi-

vidual and shared experience. Influence, on the contrary to participation or social engagement, means some kind of elicitation of change that can be visible for instance in political processes. At the intersection between participation and influence is efficient participation. Participation must be influential one way or another so that is being experienced as meaningful for the one who participates (Anttiroiko 2003, 16-17.)

According to Anttiroiko (2003, 18), *empowerment* and *inclusion* are in turn terms that are used to describe the institutional actions that aim to enhance citizens' possibilities to participate and influence. In Anttiroiko's (2003, 18) sense inclusion represents administration's efforts to increase citizen participation and empowerment on the other hand refers to actions that aim to improve possibilities of citizens to influence on the situations that affect their lives.

In the following figure are represented the field of citizen participation with contiguous concepts from Anttiroiko's (2003, 19) perspective:

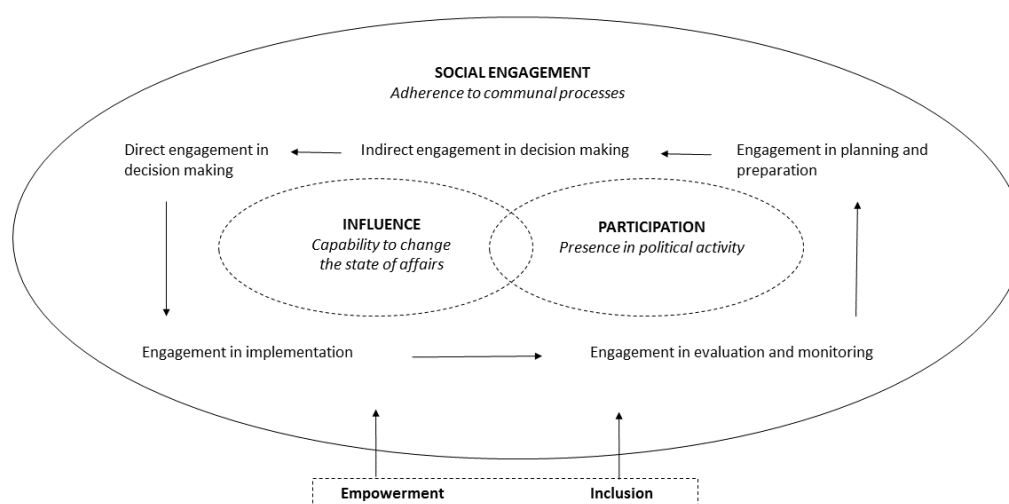


Figure 3. Field of citizen participation (adapted and translated from Anttiroiko 2003)

In his book Jan van Dijk (2012, 111-116) also discusses participation possibilities of citizens in democratic govern especially by pondering the opportunities that new media has brought for citizen participation. Van Dijk (111-112) refers these possibilities with the term *e-participation* which he sees as a part of *e-government*. According to van Dijk (2012, 111), “e-government comprises all processes of information processing, commu-

nication and transaction that pertain to the tasks of the government (the political and public administration) and that are realized by a particular application of ICT”. Thus, in van Dijk’s sense taking part in public affairs by means of ICTs to shape these affairs is e-participation. Van Dijk (2012, 112) emphasizes that participation happens in a particular phase of the institutional policy process and he describes these phases as agenda-setting, policy preparation, decision-making, policy execution and policy evaluation. In each of these phases there can be found different applications utilizing ICT to participate (van Dijk 2012, 113).

Van Dijk’s (2012, 112-115) descriptions that characterize citizens’ participation in different phases of policy process by means of ICTs resemble the cycle of engagement possibilities outlined by Anttiroiko (2003, 19) in figure 2. Both of the authors regard that citizen participation and influence is strictly connected to political processes and participation is mainly seen as participation in policy making. Although the availability and use of open data can enhance participation and influence in political sense as well (see Davies, 2010) I still regard that the approaches that consider citizen participation and empowerment only from political perspective are too limited for my examination of open data use and the possible social impacts and values originating from the use. As the data that is opened by the administration is highly versatile and considers various topics, the utilization is presumably as multi-dimensional as the nature of already opened data. Therefore the use and the benefits gained from the utilization in citizen level are not only political and data utilization can be socially valuable and empowering in other contexts as well. In order to pay attention to other contexts besides political, I will consider the utilization of open data through the perspective of empowerment theory (chapter 3) as I see it more adequate lens to view the phenomena in question.

3 EMPOWERMENT THEORY AND OPEN DATA

The literature review considering open data utilization suggests that there can be found signs that support the idea of the relation between open data use and empowerment. Although Davies (2010) is not focusing on the concept of empowerment in his study, his findings regarding civic uses of open data and democratic engagement (Table 2) can be considered from the perspective of empowerment. But what the term empowerment actually means when discussing open data use? The literature focusing on the impacts of open data use usually emphasizes the empowering potential incorporated in the data particularly as potential whereby citizens and enterprises can act more efficiently as they now have an equal access to data which is used to run the society. Despite of the frequent use of empowerment in these discussions, the term often stays rather open-ended without defined meaning. The general use of the term seems to be a wider problem extending in various fields of research and Page and Czuba (1999) note that sometimes empowerment as a term is used only to confirm the significance of certain project or goal.

In their study of the emerging impacts of open data Davies et al. (2013) also raise empowerment as one of the social impacts which is predicted to emerge along with the utilization of open data. According to Davies et al. the predicted impacts in the category of “inclusion and empowerment” are understood as follows:

“Open data will remove power imbalances that resulted from asymmetric information, and will bring new stakeholders into policy debates, giving marginalized groups a greater say in the creation and application of rules and policy” (Davies et al. 2013)

In their view open data acts as a tool or resource which will help its utilizers to associate into policy debates and creates potential to influence in societal debate. This rather optimistic and data deterministic view contains the idea of open government data being the intermediary of power between the administration and other actors within society. Actors in the lower hierarchy, like citizens, can thus use this novel resource to challenge the conventional power balance between administration and others. Changes in power balances of actors are the core of empowerment discussion but the study of Davies et al. (2013) says a little about the concept of empowerment itself and therefore it offers only minor insights to the question how these power changes originating from open data use can actually be measured. In order to understand what empowerment means in relation to

open data utilization and how the concept could be studied a better knowledge of the concept itself is needed. In this chapter I review the theoretical background of the concept and ponder how the views regarding empowerment can help the study of open data utilization particularly at individual and civic level.

3.1 Defining the concept

Like the description of the empowering possibilities of open data by Davies et al. (2013) indicates, empowerment connotes the idea of positive change in power relations of someone or something. The multidisciplinary literature about the essence of the concept shows that the definition of empowerment is, however, by no means simple or thoroughly agreed. The concept has been crucial theme for several fields of research starting from the sixties. The topic has been the subject of study for example in the fields of social psychology, education, social science, social anthropology and public health (Eklund 1999, 39.) Empowerment as a concept which relates to the transformation of power of the oppressed is also widely used in the fields of development studies and women's movement.

The origins of the word empowerment come from Latin and according to Eklund (1999, 41) it is derived from words "power" and "freedom". Thus power can be seen as a root term for empowerment and empowerment refers to the processes where individuals or groups gain power and freedom to act effectively in situations that affect their lives. Eklund (1999, 41) states that empowerment theory is based on a view of society where power as a non-material resource is differentially distributed between several groups acting in this society and by empowerment power can be transformed. According to Eklund (1999, 41), Julian Rappaport was first to introduce a definition for empowerment in the early eighties by stating "Empowerment means at enhancing the possibilities for people to control their own lives".

Since the first attempt to define the concept Rappaport and others from the field of community psychology have actively participated in the discussion of deeper understanding of the aspects regarding empowerment. In the article "Studies in Empowerment: Introduction to the Issue" Rappaport (1984) continues his view of empowerment by stating that empowerment is a process and it works as a mechanism for people to gain mastery over their lives. For some people the mechanism of empowerment can lead to a sense of

control and for others it may lead to actual control that is the practical power to influence their own lives. In the article Rappaport also acknowledges the difficulties regarding the definition of empowerment because the processes and end products of empowerment take different forms among different people and contexts. Thus, Rappaport states that empowerment is easier to define in its absence with words like powerlessness, learned helplessness or alienation (Rappaport 1984.)

Despite of the recognized difficulties to define empowerment as a single concept, Page and Czuba (1999) present the following definition based on their literature review: “Empowerment is a multi-dimensional social process that helps people gain control over their own lives.” Page and Czuba continue the definition by specifying empowerment process with the following statement:

“It is a process that fosters power (that is, the capacity to implement) in people, for use in their own lives, their communities, and in their society, by acting on issues that they define as important.”
(Page and Czuba 1999)

According to Page and Czuba (1999), their definition includes three components that remain constant in any case of empowerment despite of the context where it happens. In their view empowerment is multi-dimensional, social and a process. Multi-dimensionality of empowerment refers to different dimensions and levels within empowerment occurs. The dimensions are for instance sociological, psychological and economic among others and empowerment can take place at levels such as individual, group and community. Empowerment as a process is understood as similar to path or journey which develops on the way. The journey is also social since empowerment can only occur in relationship to others (Page and Czuba 1999.)

3.2 How different types of power help to understand empowerment?

As the definition by Page and Czuba (1999) state that empowerment is a process that fosters power the definition of power must also be considered at certain level. Seth Kreisberg (1992, 22) notes in his book “Transforming Power: Domination, Power and Empowerment” that many attempts to define empowerment fail since the discussion regarding the essence of power is ignored or forgotten. According to Kreisberg, (1992, 22) this failure is a result of the problematic nature of power itself and there has been difficulties to find conception of power that fits with the idea of empowerment. Thus, Kreisberg

(1992, 22) states that “Empowerment is a theory in search of a compatible conception of power”

The problem to connect power with empowerment originates from the predominant understanding of power which regards power as domination and the abilities to control and impose one’s will on others (Kreisberg 1992, 21). From this viewpoint power is understood as *power over*, as a zero-sum game where one’s gains of power automatically reduces the power of others (Alsop et al. 2005, 232). When discussing empowerment as a process which leads individuals or groups to gain more control over their lives, this process doesn’t necessarily reduce or impact on the power of others. Therefore the idea of power as domination or zero-sum does not associate effectively with empowerment. This notion is particularly apparent when considering open data or some other digital material as some kind of resource which embodies power. When data that is distributed and licensed openly is used by someone, the use does not limit the possibilities of others to use the same data to different means. Like Kreisberg (1992, 22) notes, alternative views regarding power are needed in order to understand better the relation between power and empowerment.

Usually these alternative types of power which can help to understand empowerment more efficiently are referred as *power with*, *power to* and *power from within*. Compared to power over, these types of power are seen as positive and regarded as productive forces. Power with refers to collaborative power and it is based on mutual support, solidarity and collaboration and thus it multiplies individual talents and knowledge. Power to is regarded as the individual potential to shape one’s life and it is seen as generative since it creates new possibilities without domination. Power from within refers to psychological and individual assets, like the capacity to imagine, have hope or believe one’s self and it is connected with self-actualization (Alsop et al. 2005, 232-233.) Table 3 by Luttrell et al. (2009) presents a helpful illustration of the typology of power in relation to empowerment from the perspective of agency and structure.

Table 3. Typology of power in relation to empowerment. Adapted from Luttrell et al. 2009.

Type of power relation	An 'agency' approach to empowerment	Transforming 'structures' for empowerment
Power Over: the ability to coerce and influence the actions and thoughts of the powerless	Changes in power relations within households and communities and at the macro level, e.g. increased role in decision making and bargaining power	Respect equal rights of others, challenge to inequality and unfair privileges
Power To: the capacity to act, to organise and change existing hierarchies	Increased skills, access and control over income and resources, and access to markets and networks	Increased skills and resources to challenge injustice and inequality faced by <i>others</i>
Power With: increased power from collective action, social mobilisation and alliance building	Organisation of the less powerful to enhance abilities to change power relations Increased participation of the less powerful	Supportive organisation of those with power to challenge injustice, inequality, discrimination and stigma
Power from Within: increased individual consciousness, self-dignity and awareness	Increased confidence and awareness of choices and rights; widened aspirations and ability to transform aspiration into action	Changes in attitudes and stereotypes; commitment to change

3.3 A multidimensional process

By following Rappaport's (1984) initial definitions, the literature usually considers empowerment as a process which leads in gains of control or power to act effectively. Along with Page and Czuba (1999), for instance Kreisberg (1992), Alsop et al. (2005) and Luttrell et al. (2009) characterize empowerment as a process or a progress. On the other hand, like Luttrell et al. (2009) note, empowerment can also be considered as an outcome, a change that is achieved. The decision how empowerment is perceived between these two viewpoints naturally affects the attempts to measure the aspects of empowerment.

From my perspective empowerment needs to be understood in both way, as a process and as an outcome, since processes always start from somewhere and lead to some kind of outcome. In order to measure and state that someone is empowered, i.e. has gained control over one's life, empowerment has to be regarded as outcome resulted from empowerment process. The recognition of starting points and outcomes is thus important when examining cases where empowerment is believed to occur. Since empowerment is rather ongoing process and it would be absurd to claim that someone can become fully empowered to control one's life, the outcomes must be considered as waypoints and associate them with the possible dimensions and levels where empowerment can occur.

According to Eklund (1999, 45), in his study Kieffer considers empowerment also as progressive process which passes through several phases. By studying emerge of individual activists in citizen organizations, Kieffer found patterns which were apparent in the

transition from sense of powerlessness to sense of sociopolitical empowerment. The transition identified by Kieffer includes four distinct phases of involvement in which individuals construct skills and insights that constitute a “fully matured attainment of participatory competence”. The phases of involvement according to Kieffer are era of entry, era of advancement, era of incorporation and finally era of commitment. This continuum of individual development is also characterized by individuals’ continuing internal “constructive dialogue” and the circular relationship of her experience and reflection. The state of empowerment requires also attainment of extensive and abiding competence in each of the following intersecting areas: development of a more positive self-concept, construction of more critical or analytical understanding of the surrounding social or political environment and cultivation of individual and collective resources for social and political action (Eklund 1999, 45.)

As the research considering empowerment is highly influenced by the field of social and community psychology (see the works of Rappaport and Zimmerman), empowerment theory usually refers to the psychological dimension of empowerment. Empowerment process is closely connected with one’s feelings about oneself and like Rappaport (1984) notes, empowerment is something that must be achieved at individual level first and can’t be directly given for people. In addition to psychological dimension empowerment is still closely linked with social and political conditions in which people live (Kreisberg 1992, 19). Also Zimmerman (1990) emphasizes the importance to expand the understanding of empowerment beyond individual level and psychological dimension of analysis as the overly individualistic standpoint can fail to consider external factors related to empowerment like environmental influences, organizational factors, or social, cultural, and political contexts. Excessively individual focus can also limit the possible research methods and designs used (Zimmerman 1990).

In his paper “Taking Aim on Empowerment Research: On the Distinction Between Individual and Psychological Conceptions” Zimmerman (1990) introduces the conception of psychological empowerment which he distinguished from the individually oriented conceptions of empowerment. According to Zimmerman psychological empowerment still refers to the individual level of analysis, but it takes into account also the ecological and cultural factors which influence empowerment. Zimmerman states that “Psychological

empowerment is contextually oriented conception of empowerment that embraces the notion of person-environment fit.” (Zimmerman 1990.) He continues by describing the goals of empowerment research:

“The goal is to understand how what goes on inside one’s head interacts with what goes on in one’s environment to enhance or inhibit one’s mastery and control over the factors that affect one’s life.” (Zimmerman 1990)

3.4 Open data utilization, an unlike approach of empowerment research

Studying the relation between open data utilization and individual and citizen empowerment processes is somehow different compared to the contexts where the concept is usually studied. Empowerment is often studied by focusing on oppressed individuals or groups but in the case of open data utilization the group under focus is not regarded oppressed per se. Like the studies of Davies’ (2010) and Okamoto’s (2016) present, the common utilizer of open data is usually a highly educated male. Therefore empowerment processes experienced by those using open data can differ from findings of other studies examining empowerment and thus results may not be directly comparable. The different focus also complicates the direct use of frameworks created for measuring empowerment in the analysis as my approach differs from contexts where empowerment is commonly studied.

Regardless of the unlike starting point, I still see that open data and its utilization associates highly with the concept of empowerment. Open data as non-scarce and free resource offers, at least at rhetorical level, myriad use possibilities that can help its utilizers to gain more control over their lives and foster power to act more effectively at individual, community and societal level. In my study I’m willing to examine the relation between empowerment and open data utilization beyond the rhetorical level. Since my primary focus is on the experiences of individuals, it can be regarded that I’m studying empowerment from individual and psychological perspective. In addition I’m still treating empowerment as a phenomenon beyond individual and psychological, one which is linked with social and political contexts. Therefore the idea of empowerment research described by Zimmerman (1990) can be seen to act as some kind of guideline for my approach.

4 STUDY AIMS AND METHODOLOGY

In this chapter I present the aims of my study and discuss the methodology that I have used to answer my research questions. In chapter 4.1 the research problem, research questions and the aims for the study are summarized. In chapter 4.2 I present and justify the chosen research approach and discuss the methods that I have used for data collection and data analysis.

4.1 Research problem and the aims for the study

As it is stated earlier, open government data can be regarded as some sort of an intermediary of power and resource that can help people to gain more control over their lives in myriad different situations. In the field of open data research it, however, seems that the social benefits originating from open data utilization are still scarcely studied. The actual knowledge concerning utilizers' experiences regarding data utilization and empowering potentials of the use remain scant, especially in the Finnish contexts. Therefore the main research question steering my study is the following:

How open data utilization empowers its utilizers in Finland?

In other words, I'm aiming to study how the utilization of raw open data can help the utilizers to gain more power and freedom to act effectively in situations that affect their lives. In order to tackle the main question accurately, I see that the discussion must be based on the analysis of the following sub-questions as I regard that these themes in question are closely related to the idea of empowerment:

- How open data is used from individual perspective and what are the purposes for utilization?
- What kind of benefits and impacts data utilization produces?
- What kind of factors enable data use among utilizers?

I believe that detailed and case-related description characterizing social benefits hailing from open data use is needed alongside with analyses examining the economic benefits. In my opinion such attempt would also profit outside academia and be valuable as more efficient infrastructure for open and participatory society is being developed further.

4.2 Research approach

The above stated research problem and the questions steering my study can be regarded qualitative as such. In addition, like empowerment theory demonstrates, processes of empowerment are highly based on individual experiences. Therefore my approach can be seen as qualitative and I make use of qualitative data collection and analysis methods in order to produce knowledge on the topic. It must be stated that I'm not aiming for comprehensive generalization about the empowering potentials embedded in open data use but instead I'm willing to produce country-specific and tentative insights to phenomenon which is still comparatively novel. Due to the novelty and the lack of earlier knowledge describing the phenomenon, it also seems challenging to conduct a relevant quantitative research examining the issue.

4.2.1 Data collection methods

Since the subjects of my study are individual open data utilizers at raw-data level and their experiences regarding data utilization, data collection was based on methods which can be used to inquire individual experiences. In my opinion the questions regarding the purposes of utilization, experiences of impacts and benefits and factors enabling empowering use can only be analyzed based on data which is gathered from informants that are familiar with open data and its utilization. Thus, I find it crucial that the informants have actual experience on the subject in the same manner like Tuomi and Sarajärvi (2003, 87-88) note when they discuss about the nature of qualitative studies.

In order to pay attention to the above mentioned notions, I separated data collection in two phases. Data collection started with open online-survey and it was followed with 4 theme interviews. The choice to use both methods, online-survey and interviews, was partly due to the qualities of my research questions but it also resulted from the anonymous nature of open data use. Since data sources can commonly be accessed freely, i.e. without any need to register to any service, the information about utilizers, their purposes and the contexts of use stays somewhat hidden. From my perspective the most suitable way to tackle the problem related to the anonymous nature of utilizers and thus possible informants was to start the data collection with online-survey that was shared openly and targeted in relevant channels. In my opinion survey also seemed as proper instrument for the so called elite sampling (Tuomi and Sarajärvi, 2003, 88-89) of informants for the

interview phase. With the survey respondents had the possibility to express their willingness to participate in theme interviews and to describe examples and projects related to their utilization of open data.

On the other hand, I also found survey being the most agile method for collecting data about the general situation regarding open data utilization in Finland. I saw it necessary to compose some sort of understanding about data utilizers, their habits to use data and their views regarding the potentials of open data before I could inquire the questions related to individual experiences of empowerment. Thus the results of the survey can be seen to form base for the structure of interviews. Altogether the survey (Appendix 1) included seven different questions. Utilizers' understanding regarding open data, its meaning and utilization potential in different contexts was explored with different propositions that respondents could associate with four-step scale (strongly, partly, weakly, not at all). Additionally, to gather more diverse data concerning these themes, respondents were also offered the possibility to describe their experiences verbally. The "Five processes of OGD use" -model identified by Davies (2010) was used as a base for exploring the ways of data use. Respondents had the possibility to select from Davies' scenarios those that describe their utilization habits the most. Besides the selection from the predefined scenarios, respondents were also able to describe their habits verbally if the processes defined by Davies' (2010) didn't correspond their experiences. The topics and contexts concerning utilization and the use examples were inquired with open ended questions.

After the survey I continued data gathering with theme interviews. According to Tuomi and Sarajärvi (2003, 77) in theme interviews the human interpretation of things and their meaning making about things are emphasized. Compared to surveys, interviews have also advantage in their flexibility, since survey questions are always fixed and prepared in advance. During an interview, the interviewer has the possibility for instance to repeat questions, change the order of the questions, correct misunderstandings and have a conversation with the informant. (Tuomi and Sarajärvi 2003, 75.) I see that aspects related to the conversation-likeness of interview are crucial when the aim is to gather data about experiences. In my opinion the individual aspect, which is emphasized part of empowerment process, can be more accurately examined with method which allows the informant and the interviewer to discuss about experiences. As the phenomenon under research is still comparatively novel, flexibility in data gathering seems particularly important.

4.2.2 Thematic structure for interviews

The following five themes formed the structure for the interviews. I created the themes in order to steer the conversation towards topics which I regarded essential and important from the perspective of my research questions. The structure is based partly on the theoretical background introduced in the chapters 2 and 3, partly on the results of the survey and partly on the qualities of my research questions and the research problem as a whole. Since all interviewees, except one, had answered the survey, the structure of each interview was also modified on the basis of informant's previous answers.

Thematic structure:

1. General interest in open data and role as utilizer

Within the first theme I was interested in exploring informant's interests in open data at general level. I aimed to examine for instance what creates the interest in open data and its utilization and in which contexts or role the informant usually utilizes data (work, hobby, active citizen). Within this theme I was also willing to discuss if the use of open data is related to certain issues, hobbies or objectives which the informant sees as meaningful for her, in order to steer the interview towards the issue.

2. Data utilization (important topics, use scenarios, examples)

Within the second theme I aimed to discuss topics and data themes which are emphasized when informant uses open data. I was interested in examining the purposes and objectives for the use and exploring the ways how the informant has utilized certain data. To further the conversation, I focused on the example cases and use scenarios that informant had described in the survey.

3. The significance of data and its availability

Under this theme my aim was to continue the discussion regarding example cases and sort out the importance and role of data in these cases. I found it important to explore if the availability of certain data creates the need for use and actions related to it or if open data is seen as utility or tool which helps forward some action or objective. I was also interested in identifying cases where the availability of certain data could help the informant to act more effectively certain situations.

4. Gained benefits and the possibilities to impact

Based on the example cases and themes that arouse in the conversation, I aimed to guide the discussion towards the benefits that the informant has gained with open data use and eventually towards her experiences regarding the possibilities to affect with open data use. Are the benefits gained only personal or has the use benefitted others as well? In which situations open data enables or can enable the informant to act more effectively for instance as citizen? In my opinion, the examination of these factors is crucial in order to assess open data utilization from the perspective of empowerment.

5. Factors enabling open data use

Within the last theme, I aimed to map factors, like skills and resources which the informant regards crucial for her data use. In addition, I was also willing to focus on the deficiencies which have weaken the informant's possibilities to use data or to benefit from certain data. I see that the mapping of these factors could further the critical conversation concerning the inclusiveness of open data use and give insights what is needed in order to enhance the wider use of open data among society.

4.2.3 Implementation of survey and interviews

After one pilot round the survey was open for answers for the whole January 2018 (4.1.2018 – 31.1.2018). In order to promote the survey and to gather relevant informants I posted an open call for answers on the public Facebook group called Finnish Open Data Ecosystem (<https://www.facebook.com/groups/fi.okfn/>). This FB-group is one of the most actively used channels for Finns to discuss the issues related to open data and therefore it seemed a fruitful source to reach relevant respondents. The group is maintained by Open Knowledge Finland organization and at the time of the survey FB-group had over 4000 members. After the publication of the survey and the call for answers, the link to the survey was also shared via Facebook and Twitter by several actors. To spread the word about my research I also contacted the administrators of the largest national and regional open data portals in Finland (avoindata.fi -portal, HRI-portal, Tampere Open Data Portal, Oulu Open Data Portal, Lounaistieto-portal, Turku Open Data Portal) and asked them to promote the survey in their communication channels.

During a one month period online the survey gathered 13 answers and six respondents expressed their interest to participate in the interviews. After I contacted these respondents via email only three of them showed real interest to participate in the interviews. The first three interviews were arranged during March 2018. Due to geographical distances and low arrangement resources for interviews, all interviews were conducted remotely by using Skype and recorder with MP3 Skype recorder -application. On April 2018 I succeeded to recruit yet another interviewee and the fourth interview was conducted in the end of April.

4.2.4 Qualitative data analysis

As the low response rate of the survey (n=13) disallowed any proper use of quantitative analysis methods, I have considered all the collected data as collection of texts that represent individual experiences regarding open data utilization in Finland and therefore approached these texts from the perspective of qualitative content analysis. Like Tuomi and Sarajärvi (2003, 93) note, content analysis can be regarded as individual method but in the wider perspective it is also considered as loose theoretical framework for the analysis of written, heard or seen content. In my approach I have utilized content analysis as a lens through which I have examined the collected texts that describe individual experiences. This lens includes some already known research on the subject, like the studies of Davies' (2010) and Okamoto (2016), and knowledge from empowerment theories but I have also aimed to produce the reasoning from the ground up. In other words, my analysis approach could be regarded as a mix of inductive and deductive reasoning that is sometimes referred as the third logic of reasoning, abductive reasoning (Tuomi and Sarajärvi, 2003, 95-97). According to Tuomi and Sarajärvi (2003, 97), the logic of abductive reasoning is based on an idea which regards that the formation of any theory is possible only when there exists some leading thought or clue connected with the observation of data. In abductive reasoning the thought process and orientation of the observer also varies between collected data and already known theoretical models (Tuomi and Sarajärvi, 2003, 99).

The leading thought for my observations and analysis has been the empowering potential embedded in open data when it is utilized from individual and citizen's perspective. Since empowerment and other social benefits originating from open data use are not studied thoroughly enough the analysis could not have been performed as completely theory-ori-

ented. The theoretical framework of empowerment and the observations made in the earlier studies examining open data utilization are still present in my analysis process but I have aimed to produce the units of analysis from the collected data itself in the same manner like in the tradition of data-driven analysis of qualitative studies (Tuomi and Sarajärvi, 2003, 97). Based on these factors, theory-bound analysis is possibly the best form that describes my approach. By theory-bound analysis Tuomi and Sarajärvi (2003, 98-99) refer to analysis process that has certain theoretical connections and theory can help to further the analysis but the analysis and the observations made are still grounded in the data. Tuomi and Sarajärvi (2003, 98) note that in theory-bound analysis the impact of the already known knowledge and theories is recognizable in the analysis process but the already known is utilized to open up new ideas and conclusions concerning the phenomenon in question. This notion is particularly apparent in my approach and I see that only with theory-bound I could produce proper analysis to answer my research questions.

My analysis in practice was based on general method of content analysis (Tuomi and Sarajärvi 2003, 105). First I disassembled and filtered the collected data and then I coded and contextualized it again to a logical whole. The disassembling and coding was initially based on the thematic structure of interviews but by following the tradition of theory-bound analysis, I used the thematic structure only as my first analysis framework and tried to let new themes and meanings arise from the data while it was being processed. In order to systemize my interpretation and advance my analysis process I have also utilized qualitative analysis techniques like thematization and typifying as my ways of orientation (Tuomi and Sarajärvi 2003, 104).

5 ANALYSIS OF THE SURVEY AND INTERVIEWS

In this chapter I present the analysis of the data gathered with online-survey and four theme interviews. The analysis aims to answer the three sub-research questions presented in the earlier chapter and thereby it creates a foundation for the discussion that concerns the main research question regarding empowering potentials of open data utilization. Chapter 5.1 is mainly based on the results of the survey and it describes respondents' affiliations with open data, their perspectives on its potential and their data utilization in general. In chapter 5.2, I examine the habits and purposes of individual open data utilization more closely by presenting a thematization of individual open data utilization. In chapter 5.3 the focus is on the analysis of respondents' experiences regarding the benefits and impacts of open data utilization and in chapter 5.4 I discuss how individual competence concerning data utilization seems to define the utilization. In the following chapters the interviewees are referred anonymously with letter-number combination like I1. In this combination the number stands for the order the interview.

5.1 Respondents' affiliations with open data and perspectives on its potential

Based on the results of the survey and theme interviews, respondents' initial interest in open data seems highly work-related. When considering the survey results, most of the respondents saw that open data is strongly related to respondent's own job and employment in the field of IT, public sector or GIS was highlighted in the results. Also the other background info that the respondents offered seems to follow the typical characterization of open data utilizer that has been presented in earlier studies (Davies 2010, Okamoto 2016). Most of the respondents were male and many informed to hold at least Master's degree in fields like natural science, geography or environmental science. General perspective regarding the potential of open data among the respondents seemed also to emphasize the technical aspects of data. According to the results, open data was commonly identified as economic and technical resource, one that enables the creation of new innovations and businesses. Most of the respondents associated strongly with the claim that open data stands for "bits" that can be utilized in software development and in most cases

open data was also regarded at least partly as raw-material which can be utilized in different business contexts.

Although the connection between open data and respondents' work role seemed strong, the results still indicate that open data is seen to be linked with other aspects of life as well. According to the survey, the respondents regarded open data also as social utility which has significance from individual and citizen's standpoint. Each respondent associated either strongly or partly with the claim "Access to administrative data is one of my privileges" and most of the respondents experienced that open data utilization helps the supervision and improvement of the actions of administration. Open data was also linked to respondents' hobbies or other leisure-time activities among the results and some of the respondents saw strong relation between open data use and their acts as "active citizens".

Theme interviews strengthened the observation that in addition to work related use, open data is also utilized from individual standpoint. Two interviewees saw that their utilization of open data was at that time happening only in their leisure-time and they considered their data utilization as one of their hobbies. The other two described that in their current job their work is more or less related to open data, but still they told that they have utilized open data from individual perspective as well. In these cases where the work-relation was stronger, individual and leisure-time utilization seemed to follow to some extent their work utilization in the light of data topics and utilization techniques. All of the interviewees still experienced that their initial awareness and knowledge regarding open data and its use is somehow connected with their current or previous jobs and their educational background.

In addition to the many-sided views regarding the utilization potential of open data, the actual data use of the respondents seemed versatile. Each survey respondent reported having used at least two different Finnish sources of raw open data and also foreign sources were utilized. The most used sources among respondents were Statistics Finland, avoindata.fi -portal, HRI-portal, Finnish Meteorological Institute and National Land Survey of Finland. Although the "use" of certain source in its simplest meaning could only denote the visit in web service, it seems that each respondent had also utilized open data somehow, since each respondent had picked at least one of the use-scenarios defined by Davies (2010) when describing their habits of use. The most popular scenarios of data use among respondents were "data to information" and "data to fact". The combination

of different datasets and data visualizations were highlighted as popular ways to utilize data in the free form descriptions.

Overall the group of utilizers who answered the survey can be regarded, at least to some extent, as experienced and skilled on behalf of their data use. Many respondents reported that they had utilized data according to at least three different scenarios identified by Davies (2010) and some of the respondents had also either created a service based on open data or utilized open data in some already existing service.

5.2 Themes of individual utilization

The results of the survey offered only very general knowledge concerning respondents' utilization habits but the data gathered with interviews helped me to form a more comprehensive look on how open data is being utilized individually. Based on the combination of data gathered with survey and interviews, the actual utilization of data among the informants seemed to associate with different themes that describe the informants' utilization habits and aims of data use. I have identified these themes as *data to understand*, *optimizing with data* and *data as utility for skill development and self-expression*. The essence of each theme with some demonstrative example scenarios is presented in the following chapters. These themes are created especially to describe informants' open data utilization happening in leisure-time, i.e. the example cases of data utilization in which the themes are based are not work-related. As the thematization concerns only individual and leisure-time utilization habits, my aim is to create a framework that could be used when open data utilization from individual and citizen perspective is explored further.

5.2.1 Data to understand

According to the interviews and the survey, one of the most emphasized themes of data utilization among the informants was the utilization of available open data to understand certain issue, case of interest or things in general. Based on the example cases among my sample it seems that data can be utilized in order to understand highly specific issues and topics but also more mundane topics. When the aim is to understand some specific issue with the help of data, like the future energy consumption scenarios in Finland (I3), it seems that in order to understand, the utilization process requires more high-level analysis techniques and skills to work with data. On the other hand, data can help to understand

with lesser efforts needed, for instance in cases where understanding is formed by browsing data with the help of applications like Excel.

In free-form descriptions of the survey open data was characterized as a source for social understanding, one which allows individual reasoning and creation of up-to-date understanding about the state of the surrounding society. An example given by one respondent presents well how open data can be utilized in order to improve individual understanding about society and politics:

“I have for instance made tool for myself that I can use to get my hands into data concerning the election financing of the last parliamentary elections in order to make different comparisons between parties and areas but also to explore which actors are financing and who.” (Survey respondent)

Based on the utilization examples highlighted in the interviews it seems that open data can particularly help to understand things that are related to the current state of nature and climate. For instance I4 had made visualizations about the history of hurricanes around the Atlantic Ocean and the melting of the Arctic ice and I3, among other environment data related use, had made a program which is capable to model in real-time the spreading of toxic fumes and gases in case of some accident. This savor to understand topics related to nature and climate with the help of open data can be due to the general interest in environmentalism which was highlighted among interviewees. For instance I3 regarded that environmentalism is important topic for him. On the other hand I3 told that at the moment he was particularly enthusiastic to work with real-time data and in the Finnish context lot of weather and climate related data are shared in real-time. I4 also emphasized that he is interested in data which is plenty available and has good quality.

The results indicate that utilization of open data can help people to understand better certain issues or cases of interest but it seems that utilization can also enable individual monitoring and supervision of things happening around. In the survey most of the respondents experienced that open data embodies the potential to improve the supervision of administration as most of the respondents associated strongly with the claim “With open data the actions of administration can be supervised and improved”. Along with the beliefs in supervision potential, open data seems to actually enable possibilities for individual monitoring and supervision if considering the utilization examples of I3. I3 for instance told that he is running a program in his server which tracks and records in real-time the ship

traffic in the Baltic Sea to monitor the exceptions and accidents happening. Also the program that I3 has developed to model the spreading of toxics in the environment seems an example of data utilization as monitoring and supervision. Also the descriptions by I4 about his aims concerning his data utilization point at the monitoring and supervision potentials of open data:

“When I visualize data, I try to visualize it in the right way and sometimes I may try to dig up some truths from the data” (I4)

I4 continues by stating that his data utilization can be regarded as “one man battle against fake-news”.

5.2.2 Optimizing with data

In addition to utilization theme identified as data to understand another slightly different theme of open data utilization was apparent as informants’ described their data utilization. I have identified this second theme as optimizing with data and it represents a form of utilization where certain data available can help or is used to enhance different individual and social actions. Additionally, the optimization can relate to cases where utilizer has individually produced or modified open data in order to enhance something. Based on the examples reported by interviewees, optimization can concern either utilizer’s own private life, public services or utilizer has enhanced the data use of others.

The case of garden planning described by I1 is a demonstrative example of how open data can help to optimize things or actions related to utilizer’s own private life. According to I1, she has made a map-based construction plan for her garden that was based on aerial images and open property data. The aim of the utilization was to enhance the whole construction process by calculating the exact material needed for the construction. When contrasting on her description, the optimization seems to concern particularly her money spending related to material expenses but with the help of easily accessible data she also managed to enhance her time spending and lower the efforts needed for the project.

Another optimization related utilization habit was also apparent as I1 described her use of open data. She had for instance improved the functioning of the trip planner service of her local public transportation by creating new geographic information in the service. I1 also mentions that sometimes she adds her own sightings of certain species in the database

of Finnish Biodiversity Info Faculty via online-service called Laji.fi. In both of these example cases the action of optimization is not based on available data and it is not only concerning utilizer's private life but rather the aim is to enhance some existing public service by creating new and more accurate data for the service.

In addition to the above dimensions optimization yet another dimension was apparent as interviewees described their experiences related to open data. This dimension of optimization concerns the enhancement of possibilities of others to utilize raw open data more effectively. For instance I3 reported that a while he ran a program in his own server which converted the newest data opened by Finnish Meteorological Institute from xml-form to more simple data form. According to I3, the aim of this action was to help others to utilize the data opened by FMI more easily as he had noticed that people had difficulties to make use of the original data-form brought by FMI. I1 and I3 also reported that they have done some sort of mentoring concerning open data utilization in their leisure-time. I1 told that she is running a local map-making club which focuses on sharing knowledge and utilization possibilities of maps and open GIS and I3 has tried to share his knowledge regarding open data utilization for other members of The Finnish Association for Nature Conservation.

5.2.3 Data as utility for skill development and self-expression

For the forms of open data utilization described above the substance and the knowledge that can be extracted from the data seems to form the basis for the use. According to the interviewees and respondents' comments in the survey, open data, however, is also used without the keen interest in the substance of the data. Based on informants' descriptions, data is often being used from experimental basis and data can act as a tool or utility for the development of skills related for instance to programming. In the following I4 describes how open data utilization is related to his learning of programming:

“When I for example want to experiment something with Python, then I try to find some relevant data for the experiment” (I4)

Based on the multiple utilization examples reported by I3, it seems that the possibilities to experiment and “play” with open data are in big role in his interest in data. He also mentions that he's aim is to “learn something new every day” and by experimenting with open data he can develop his technical skills.

Besides being regarded as tool for developing technical skills, open data was also regarded as utility for art making and self-expression among the informants. The possibility for self-expression through the creation of data visualizations was particularly typical for I4 and he mentions that in many cases the visual appearance of the end-result is more important than the actual knowledge that the visualization process can produce. I4 also believes that his visualization hobby includes some kind of artistic aspect as he enjoys if he manages to break the boundaries of traditional information visualization by doing visualizations “wrongly”. In addition to I4’s self-expressive utilization with open data, I1 also mentions that she is interested in aesthetic visualizations based on data and she has already done some posters based on road-data for decoration purposes.

In the following table all of the identified themes of individual open data utilization are presented:

Table 4. Themes of individual open data utilization

Themes of individual open data utilization	
Data to understand	Individual monitoring and supervision
Optimizing with data	Optimization of own life’s actions Optimization of public services Enhancement of data use of others
Data as utility	Skill development with data Self-expression with data

5.2.4 Goal-oriented vs. experimental data use

At more general level the individual open data utilization among my results seems to toggle between experimental and goal-oriented use. Utilization habits identified under the theme of optimizing with data indicate more goal-oriented and problem solving type of data utilization and the experiences of I1 suggest a point of view that regards that the available data can offer solutions to some already existing problems. On the other hand, the results also indicate that data utilization can be more experimental by nature, a fact that was also highlighted in Davies’ (2010) study. In the interviews I3 and I4 emphasized their desire to try different things with data and according to them, “playing” with data

does not always need to produce new insights or answers to some particular problem. According to I3 and I4, one of the aspects that motivate them to experiment with data is still the unexpected insights and results that these experiments sometimes produce. The different themes of individual open data utilization can thus be demonstrated on the continuum between goal-oriented and experimental data use:

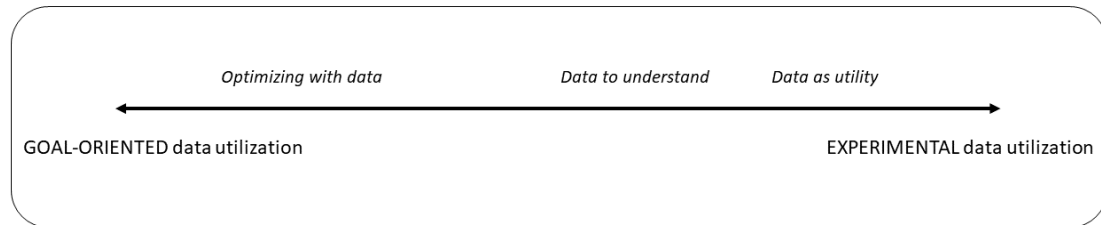


Figure 4. A continuum of individual open data utilization habits

In my view the utilization cases identified under the theme of optimizing with data are located at the goal-oriented end of the continuum as the descriptions of the informants indicate that there can be found certain goals or purposes behind the use. On the other hand, the data utilization identified under the themes of data to understand and data as utility seem to locate from my perspective on the experimental end of the continuum. Although there always exists aims and purposes behind every use, I still consider that the means of data utilization under these themes appear more experimental.

The thematization and the above continuum that I have outlined based on my results must not be regarded as exclusive or strictly defined. I regard these more like initial efforts to illustrate the diverse nature of the purposes and possibilities concerning open data use from individual perspective. Taking into account the sparsity of informants in my sample, the accuracy and feasibility of the thematization must also be considered critically. Thus, the thematization can be seen to act as a preliminary framework towards wider understanding of the phenomenon. Although the themes of utilization that I have identified represent different approaches to data utilization, it must be noted that the utilizers can always have multiple purposes and open data utilization of the informants seem to fit under many themes.

5.2.5 The thematization in comparison with earlier research

The individual utilization habits and the examples cases among my results resemble for instance the OGD use identified by Davies (2010) and Okamoto's (2016) results regarding New York City's open data utilization. The studies of Davies' and Okamoto's, however, are not examining open data utilization only from individual and citizen's perspective and thus their results incorporate examples of data utilization also from business perspective which complicates the direct comparison. The approaches of these studies seem also a bit different as data utilization is being examined more as a technical process towards some tangible outcome. In my approach though I'm more interested in exploring the outcomes of data utilization as themes that individuals gain from the process without focusing too much on the technical side of the processes. I see that it is necessary to examine the outcomes of data utilization in a wider perspective, one that goes further than the scope of created formats like interfaces or services, in order to grasp the benefits and impacts related to data utilization. This then enables more opportunities for the discussion of empowering potentials of open data.

Davies' (2010) framework of OGD use (Table 1) still seems to represent well the common technical processes of open data utilization. Based on my results, all of the five processes appeared in the example cases of data utilization described by the informants. I also see that the thematization that I have identified could work as a higher-level extensions for Davies' model. For instance under the theme of data to understand data was utilized according to processes like data-to-fact, data-to-information and data-to-interface, but besides only producing interfaces or visualizations, utilizers seemed to aim at better understanding of something with the help of these processes. Also utilization examples identified under the other themes included processes described by Davies. The following figure presents my idea of the extended model of Davies':

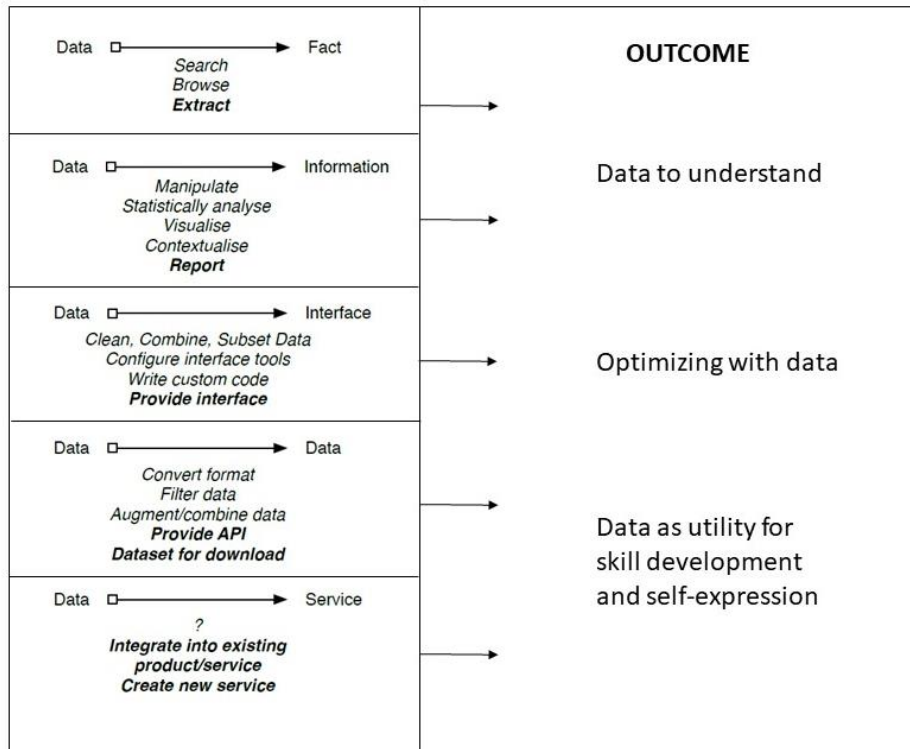


Figure 5. Themes of individual open data utilization embedded in the framework by Davies (2010)

5.3 Experiences regarding benefits and impacts of open data utilization

The results of the survey and interviews indicate that utilization of open data among informants was relatively diverse but informants' experiences regarding actual benefits of data utilization for instance as citizen seemed scarce. In some cases open data utilization has benefitted the utilizer economically. For instance I1 describes that free and open accesses to different kind of data has made possible her current job and thus affected her living. She also connects the example regarding garden planning to economic benefits that are due to utilization of open data. More common experience related to benefits of open data among the informants was that data enables nice and interesting hobbies. For instance I4 contrasts his efforts with data visualizations to carpentry and sees that crafting with data is good counterbalance for his daily work. I3 sees that his interest in programming, number crunching and statistical analyses inspires him to play with open data and hence he can learn and produce understanding from data, but also develop his skills with data. Based on the interview, I3 seems very active to individually develop different applications and analyses based on open data but despite of these activities, he has not benefitted economically nor is interested in commercializing his applications.

When pondering the question regarding the benefits of open data utilization from citizen's perspective, I1 and I2 saw that in order to benefit from open data as citizen, raw data must first be refined in more accessible and comprehensible form or offered for use through applications. I1 summarizes her opinion by stating:

“For citizens the major benefits of open data come from the applications that some actors are offering” (I1)

I2 believes that one of these applications that could actually be beneficial from citizens' perspective is a more efficient and straightforward system for the publication of official documents produced by public officers. According to I2, open system for browsing official documents could then simplify citizens' access to information by overriding the current practice of making requests of information.

The general opinion among the informants also concerned the idea that the more quantity of diverse data is published by *de facto*, the more utilization and benefits the publication would produce from the perspective of citizens and enterprises. This idea regarding the overall growth of open data utilization as the quantity of data grows resembles the long tail phenomenon introduced by Chris Anderson (2006). In addition to only believing that utilization and benefits would increase with the publication of diverse data, I2 during his work project observed actual niche for unusual data. He describes that after they had digitized and opened the material of some provincial leaflet archive they discovered that a local association interested in old tractors was increasingly using the newly digitized material. I2 regards this event as a good example of the difficulties to predict which kind of data would be beneficial from the perspective of utilizers and thus it reflects the difficulties regarding the prioritization of data to open.

5.3.1 Social influence as the by-product of experimental utilization

The belief in increased possibilities of citizens to influence on societal issues with the help of open data is always included in the discussion of potentials and benefits of open data. The survey results and interviews indicate that open data can be utilized for purposes that improve individual understanding about the society but the experiences of informants do not directly refer to enhanced possibilities to influence as citizen with the help of open data. The less active open data utilizers I1 and I2 experienced that their potential to influ-

ence for instance on local or city level hasn't enhanced with the help of open data utilization. I1 thinks that for her the only ways to impact on things with open data have been the cases where she has modified or created data by herself.

The experiences of the more active utilizers, I3 and I4, in turn seem a bit different. Based on their descriptions, their experimentally oriented utilization with open data appears to, at least to some extent, enhanced their possibilities to influence on certain topics or issues as citizens. In these example cases the influence on the issues could be seen as indirect and according to I3 and I4, social influence is not the end in itself of their data utilization. Thus, their possibility to influence with the help of open data seems more like a possible by-product of their hobbyist and experimental utilization. In the case of I3, his applications that focus on individual monitoring and supervision of topics like the ship traffic around the Baltic Sea or the spread of fumes and toxics seem to have aroused the interest of others and thereby these examples can be regarded as influential. According to I3, with the help of his ship traffic application running in his server he once discovered some unusual events happening in the Finnish coast that proved to originate from oil spill and response acts for the accident. The oil spill accident was not officially reported and based on I3's monitoring application a reporter intrigued to find out more and write about the accident. Also the development of application that models the spread of toxics in the air gained the attention of public officials and according to I3, he was invited to present the application for the Finnish Radiation and Nuclear Safety Authority during a safety training. I3 also mentions that some foreign officials had later implemented an application for safety monitoring that was based on his application.

For I4 the sharing of his data visualizations seems to link up with his data utilization and by sharing he sees the possibility to influence on others as well. According to I4, by sharing his visualizations for instance via Twitter he at least wants to believe that the visualizations and the information embedded in these could actually have an impact on someone's beliefs. Although I4 emphasizes that influencing others thoughts is not the basis for his data utilization, the hobbyist approach and sharing of own works still seems somewhat influential. Different media like Helsingin Sanomat and Yle have shown interest on I4's visualizations and used some of his visualizations on their articles. According to I4, he is also currently working as freelancer for these companies as a result of his data visualization hobby. I4 notes that in the most cases his visualizations have been more like

supplements for some already designed story or article, but I4 also regards that in some cases an article has been based on data and visualization ideas discovered by him.

In addition to above examples describing the indirect influencing with open data, some examples of data utilization among my sample suggest that open data can also help citizens when the aim is to impact more directly on issues at city or local level. I3 for instance describes that by utilizing open data concerning noise reports of some area, he has been able to make use of “noting” during zoning processes of that area and thereby affected on the validity of the zoning process. An example case described by one respondent in the survey also suggests how open data use can improve the possibilities to impact on issues on more local level. In this case a map visualization based on open data was utilized to gain and spread understanding about the potential of solar energy use around one neighborhood in Helsinki. According to the respondent, this interactive visualization acted as crucial part of the process to engage the neighborhood condominiums to purchase and utilize solar energy.

In my view the above examples indicate that open data and its utilization can particularly help to strengthen the utilizer’s individual voice as a citizen. For instance in the case of I4, by creating visualizations from data and sharing them, the issues and topics that matter can be presented for others in interesting form and backed with “facts”. As the visualizations of I4 have also gained wide attention of others and different media have used his works, the more his individual voice can be seen to have strengthened. In the case of I3, his interest to monitor and supervise things with the help of open data can in turn be regarded to have resulted in strengthened individual voice, particularly towards public officers and administration. As “facts” can be rooted in data that is available for citizens, the data can also be seen to strengthen citizen’s voice and possibilities to notify or revise the faults happening in regulatory processes.

Although the examples of enhanced opportunities for societal influence with open data were only far apart among my results and influencing with open data seems not to be the purpose for utilization, the theoretical framework of participation by Anttiroiko (2003, 19) can help to identify the types of influence in these cases more clearly. As Anttiroiko (2003, 17) defines influence as some kind of elicitation of change that visible, the only example of actual influence due to the utilization of open data seems to be the case of I3 where open data utilization has helped him to revise the faulty zoning process. The other

experiences and actions of informants seem to fit more with the definition of participation by Anttiroiko (2003, 14). If participation means actions to be involved with other people in social processes or the presence in political activity, in my view the sharing of visualizations by I4 or the monitoring activities by I3 are exemplifying how open data utilization can enhance the possibilities to participate.

Anttiroiko's (2003, 19) model of participation, however, stays at very general level and thus it is not that practical tool for closer examination. On the other hand, Davies (2010) in his study also discusses how open data utilization can enhance the opportunities for participation by outlining different civic uses of OGD (Table 2). The model by Davies thus seems more adequate for the examination of the enhanced participation possibilities that appeared in my results. In his model, Davies has identified three types of participation that he regards that the utilization of open data can enhance. Under each type of participation Davies has identified different opportunities enabled by open data utilization that can enhance the participation of that type. When considering these opportunities described by Davies in the light of my results, it seems that many of these are also apparent in my results. For instance based on the utilization examples described by I3 and I4, it seems that open data supports scrutiny and data can increase the understanding of specific issues, but also help the utilizer to inform others on these issues. In Davies' sense opportunities like scrutiny or informing of others on specific issues are seen to enhance political participation but in the light of my results, the participation with these opportunities does not seem directly political.

My results are also supporting the view that utilizers can participate in co-production with the help of open data. For instance the utilization examples presented by I1 indicate that with open data she has opportunities to participate in co-production of services and information. In the case of I3, it can be seen that by developing the toxic calculation application that is based on open data, he has had opportunity to participate in the development of public services.

5.4 Individual competence defines data utilization

At the simplest raw open data can be published as static downloadable files in formats that can be opened and processed with widely known office software like Excel. In addition to these static downloadable files, open data is yet published in multiple formats

depending on the type and field of the data and sometimes data is only accessible through APIs. To access data via APIs requires some sort of programming skills and thus the person who wants to utilize data through APIs need to have these. The experience and skills related to programming were also notable factors as informants pondered their possibilities to utilize data. The comparison between the experiences of the interviewees indicates that the more experience and skills regarding programming the utilizer has, the more versatile and complex his utilization habits of open data seem to be. For instance I3 describes that he has been programming since 1970s and he regards that his programming skills are on quite high level. Among the interviewees his utilization of open data seems also the most versatile and complex, including for instance application development and high-level data analyses.

The comment by I4 presents well how programming skills are in close relation with the possibilities to use open data:

“At the point when I was able to write the kind of program by myself which I can use to get access to data through certain API, my opportunities to utilize open data multiplied.” (I4)

I4 also regards that his skills related to programming and the use of APIs has allowed his own utilization become more format-free. On the other hand, according to I1, the little knowledge concerning programming and utilization of different APIs has narrowed her possibilities to use open data. Therefore, she often utilizes only geographic information as she regards it as her core know-how and she has technical skills to utilize it efficiently. In the following comment I1 describes the difficulties that are related to the use APIs:

“These REST APIs and the others, these are not meant to be used by ordinary people or normal utilizers” (I1)

According to the interviewees, the substance knowledge concerning data and its origins is also crucial when the aim is to do something with the data. I1 for instance emphasizes that in order to effectively utilize certain type of geographic information, she must be familiar with the data at meta-level as this knowledge eventually defines what can be done with the data. According to I1, her substance knowledge regarding geographic information is mainly evolved as a result of her education and work-career. When pondering the examples of utilization and data topics apparent in use of other interviewees, the topics manifested in their use and their utilization techniques seem as well be somehow related to their work and educational background. These facts thus indicate that in order

to use certain data efficiently, the utilizer must be familiar with the data also at meta-level.

5.4.1 Factors hampering open data utilization

Besides the lack of programming skills and the lack of leisure time, couple of other factors were also apparent as interviewees pondered factors that hamper their open data utilization. One of these is related to the lack of standardization regarding data publication and data quality. For instance I1 dreams of the utilization possibilities of recreational data of the whole Finland but she sees that the problem to find and use it efficiently is at the moment due to the fragmentary nature of its publication. According to I1, there are lots of ongoing projects that produce and publish recreational data openly and different associations are also independently producing data to their own needs, but the publishing is not managed by any official body and thus it lacks standardization and co-operation. This then complicates the efforts of the possible utilizers to find and use the data, although the data can already exist openly in some location over the internet.

The experiences of I3 and I4 suggest that there are also shortages concerning quality of open data and the lack of standardized quality of data can limit the use of data. I3 for instance refers to his urge to use air quality data for different purposes but he sees that the quality of open air quality data does not meet his needs. I4 has also paid attention to data quality, both in good and bad, and has noticed that his general image regarding the reliability and continuity of several APIs is somewhat dubious. He regards that most of the data published for instance in Finland is sufficient enough for his hobbyist use but when he ponders the possibilities to create some productive application or service based on these data interfaces he becomes more suspicious. The following comment by I4 sums up his views regarding the functioning and quality of available APIs:

“...in my opinion there is no certainty at the moment that the API will be functioning tomorrow also and if the data is incorrect, who is responsible for it?” (I4)

An example case described by I1 also indicates how the work role and role as citizen can sometimes be mixed in the world of open data thus complicate the efforts to utilize data. I1 refers to a situation where she was requesting official information as citizen but according to I1, this request caused “a lot of controversy” as the officials linked I1’s request to her work role and the organization she was working at the moment.

Altogether the results of the survey, the interviews and the above presented analysis regarding the skills and other factors needed for open data utilization indicate that for non-data-competent individual citizen, the possibilities to efficiently utilize open data are scant. According to the interviewees, besides the lack competency to utilize data citizens' general knowledge regarding open data and its utilization seems to be tenuous. For instance I2 thinks that "mainly those who are actively dealing with open data know something about it" and he believes that open data as a general topic does not emerge in the discussions of citizens. Also I3 regards that citizens' knowledge concerning open data is limited. I4 believes that the data publishers should also improve their communication regarding open data in order that the knowledge regarding open data could expand.

6 DISCUSSION & CONCLUSIONS

In the following chapter I discuss the results of my analysis in the light of my main research question, *how open data utilization empowers its utilizers in Finland*, and outline the key findings concerning empowering potentials of open data. In chapter 6.2 I evaluate the success and the limitations of my study in general and present views for further research regarding the social impacts of open data utilization.

6.1 Empowering potentials of open data utilization

The open data movement in general and the recent acts by governments to open their data for the use of enterprises and individual citizens can as a whole be regarded as a process of empowerment. By opening the formerly exclusive access to different data and information administration can be seen in the light of typology of power (see Table 3) to relinquish some of their possessed power over citizens and transform their power towards the idea of power with. This empowerment process by administration is also believed to empower citizens to act more efficiently in the society. The prediction presented by Davies et al. (2013) in their study of emerging impacts of open data describes well some of the assumed outcomes of data utilization from societal perspective:

“Open data will remove power imbalances that resulted from asymmetric information, and will bring new stakeholders into policy debates, giving marginalized groups a greater say in the creation and application of rules and policy” (Davies et al. 2013)

In the fields of social and community psychology is, however, commonly regarded that empowerment is first and the foremost individual process that fosters power in people for use in their own lives, their communities and in their society. According to Rappaport (1984), empowerment as such can't thus be directly given for people from outside as it must first be achieved at the individual level. Therefore the aim of my study has been to question the empowering potential of open data as given and by focusing on individual experiences and utilization cases of open data explore how open data utilization can actually empower the utilizers as individuals and citizens.

When considering the results that I presented in chapter 5, it can be argued that informants' experiences and example cases of data utilization among my sample do not directly

support the prediction by Davies et al. (2013) about the potential of open data to empower. At least if the prediction is pondered from citizens' perspective. The findings related to citizen influence or effective participation actualizing with the help of open data utilization were limited among my results. Based on my findings, open data and its utilization can be seen to strengthen utilizers' individual voice as citizens in certain issues in question and enhance indirect opportunities for participation but examples of utilizers' enhanced participation opportunities in the creation and application of rules and policy with the help of open data were not apparent in my results. When reflecting these observations with the idea of empowerment and typology of power (Luttrell et al. 2009), it seems that at the moment open data utilization does not significantly enhance utilizers' capacities to act or to organize and change existing hierarchies as citizens and thus does not foster citizens' power to act.

On the other hand, if the results are discussed from individual perspective the diversity of data utilization and the different utilization themes identified indicate that open data can still be seen as utility which helps utilizers to gain more power and freedom to act effectively in situations at individual level. From the perspective of typology of power, it seems that open data utilization can particularly foster utilizer's power from within. As power from within, according to Luttrell et al. (2009), is regarded as "increased individual consciousness, self-dignity and awareness" the utilization habits discussed under the theme of data to understand can in this light be seen to increase utilizer's individual awareness and consciousness concerning the field of data that is utilized. Based on my findings, open data can for instance help to increase consciousness and awareness concerning the state of environment and climate as data utilization related to these fields was emphasized among my sample. The results also indicate that utilization can foster utilizer's power to act in situations that concern her own life. According to my findings, it seems that open data can for instance help to optimize individual actions, like time use and money spending, and data is also experienced to expand utilizers' capacities to develop her own skills and express oneself.

As open data utilization seems to empower the utilizers particularly by expanding their individual consciousness and awareness (power from within), it can be argued by following Luttrell et al. (2009) that the increased power from within can also foster utilizer's capacities to act as citizen in the future. According to Luttrell et al. (2009) the increase of

consciousness and awareness also widens aspirations and the ability to transform aspiration into action. Thus, when data is utilized to understand better for instance the state of environment and climate or the funding behind political elections, the increase of awareness regarding these issues can also foster individual confidence to act on these issues. But like my results indicate, the availability and access to data alone is not enough to foster utilizers' capacities to act and influence as citizen. Therefore I see that other dimensions and opportunities for citizen participation and influence in the scope of e-democracy must be considered more closely in parallel with open data utilization so that the empowering potentials of open data for citizens can be actualized more efficiently in the future.

The discussion regarding empowering potentials and impacts of open data must also in my view proceed hand in hand with the discussion of the inclusiveness of the actualization possibilities of these potentials. For instance Gurstein (2011) has already paid attention to the problem of "empowerment of empowered" and he notices that although open data is commonly regarded to empower everyone, the potentials of open data can only be implemented by "those who are already reasonably well provided for technologically and with other resources". This claim by Gurstein (2011) seems also to hold true in the light of my results. The experience related to programming and skills to access data through APIs appeared as particular factors that seemed to divide the possibilities of utilizers to utilize data effectively among my sample. Based on my results, the initial interest in open data and knowledge regarding its utilization possibilities were also highly affected by utilizer's own work or educational background and thus the appearance of the common man as utilizer of open data was absent among my sample.

In my opinion more attention must be paid to this observable divide between those already empowered to harvest the potentials of open data and those with lesser resources and knowledge if open data is to be considered as empowering in the future. Many studies have already shown that open data at raw-data level is utilized by group of people that is somewhat homogenous on behalf of their education and work background and those interested to utilize data seem to already have skills to work with data. As raw data can be regarded to form the basis from which knowledge and wisdom are derived, the group of people that can utilize raw data can be seen to strengthen their position as intermediaries of knowledge and wisdom. I am not considering this direction merely as negative but like Gurstein (2011) in his article refers, there already exists events where the potential of

open data as positive power and productive force has turned on his head and those with resources have utilized data as coercive force against the powerless. Therefore I see that in the future it is important to focus on the identification of different utilizer groups of open data and consider more closely how those with lesser resources and knowledge regarding open data utilization can be invited to share the benefits of data utilization.

6.2 The limitations of my study and suggestions for further research

This study has been exploratory in its nature and the results regarding individual open data utilization and the empowering potentials of utilization must primarily be considered as tentative. With my study I have aimed to produce new observations concerning phenomenon which is comparably novel and therefore scarcely studied. In my opinion there is also recognizable gap in research, particularly in Finnish context, as impacts of open data are mostly examined from economic and business perspectives instead of focusing on the grass roots level social impacts of open data utilization. Therefore with this study I have aimed to fill a part of this gap and the results can be seen to act as initiative for the wider research of social impacts related to open data utilization in Finland.

In my opinion my preference to approach the question regarding the empowering potentials of open data utilization from qualitative perspective has been rather fruitful. By combining online-survey and in-depth theme interviews I have been able to collect experiences and utilization examples that express the diverse nature of individual open data utilization in Finland. With the help of abductive reasoning and theory-bound context analysis I have managed to produce a thematization that represents the habits of individual open data utilization in Finland. The thematization and the analysis concerning the experiences of benefits and impacts of open data utilization has also enabled me to discuss the empowering potentials of open data utilization in a bigger picture.

I believe that my results and the discussion concerning the empowering potentials of open data can be beneficial for the further research but also worth a look for actors who are involved in the development of the infrastructure for open and participatory society. The results must still not be taken for granted and critical attention have to be paid particularly to the generalization possibilities of the results as the amount of informants in my sample was very low. Thus, I see that the small sample of informants and the scarcity of my data

are the biggest weakness of my study. It seems that with the open online-survey I managed to reach only very few Finnish open data utilizers and thus the amount of informants in the interview phase was also scant.

From my perspective the difficulties to recruit relevant informants, i.e. actual utilizers of open data, for the study is partly due to the anonymous nature of open data use. As the utilizers are unknown it is also difficult to identify relevant channels where the possible utilizers could be reached. I also see that as I have observed the phenomenon from outside the core network of open data “activists” in Finland my field knowledge regarding the actual utilization can be somewhat tenuous. On the other hand, the weaknesses regarding my data can be due to my approach. As I aimed to examine open data utilization from wide perspective without limiting my focus on some particular group of utilizers or to some already known utilization scenario, the wide approach caused difficulties for me to recognize the possible utilizers and thus the possible informants.

Based on the knowledge that this study has produced, I see that in the future would be fruitful to focus on the better recognition of actual utilizers of open data so that the experiences regarding empowering potentials of open data could also be examined more generally. If utilizers are recognized more efficiently, the initial knowledge regarding the phenomenon could be tested and expanded for instance with quantitative approaches. In addition to examining open data utilization only at raw data level, I see that the further research could also focus on end-level utilizers of open data in order to reach the perspective of the “common man” concerning the potentials of open data. This could for instance be achieved by studying some already built applications that are based on open data. On the other hand, studies that approach the questions of empowerment and inclusiveness of open data from design perspective could also produce novel insights. In other words, the research problem could be turned upside down by asking how power and freedom of certain group of people to act more efficiently could be fostered with the help of open data. The idea would thus be to define more closely the possible empowering potentials of open data for different groups of people beforehand and then ponder how these potentials could be actualized and implemented.

REFERENCES

- Alsop, R. Bertelsen, M. Holland, J. 2006. Empowerment in Practice. From Analysis to Implementation. World Bank.
- Anderson, C. 2006. The Long Tail: Why the Future of Business Is Selling Less of More. Hyperion.
- Anttiroiko, A.-V. 2003. Kansalaisten osallistuminen, osallisuus ja vaikuttaminen tietoyhteiskunnassa. In: Bäcklund, P. (Eds.). Tietoyhteiskunnan osallistuva kansalainen, p. 11-31. Tapaus Nettimaunula. Helsingin kaupungin tietokeskus. Tutkimuksia 5-2003.
- Davies, T. 2010. Open data, democracy and public sector reform. A look at open government data use from data.gov.uk. Based on an MSc Dissertation submitted for examination in Social Science of the Internet, at the University of Oxford.
- Davies, T. Perini, F. Alonso, J. 2013. Researching the emerging impacts of open data. ODDC Working Papers #1.
- Eklund, L. 1999. From citizen participation towards community empowerment: an analysis on health promotion from citizen perspective. Academic Dissertation. Faculty of Medicine. University of Tampere.
- Gurstein, M. 2011. Open data: Empowering the empowered or effective data use for everyone. First Monday. 02/2011. <http://firstmonday.org/article/view/3316/2764> (Retrieved 4.11.2018)
- Janssen, M. Charalabidis, Y. Zuiderwijk, A. 2012. Benefits, Adoption Barriers and Myths of Open Data and Open Government. Information Systems Management. 09/2012.
- Johnson, J. 2014. From open data to information justice. Ethics and Information Technology. 12/2014.
- Kassen, M. 2013. A promising phenomenon of open data: A case study of the Chicago open data project. Government Information Quarterly. Volume 30, Issue 4.
- Kitchin, R. 2014. The data revolution. Big data, open data, data infrastructures & their consequences. Sage Publications.

Koski, H. 2015. Avoimen tiedon vaikuttavuus -esitutkimus. Elinkeinoelämän tutkimuslaitos. Ministry of Finance publications. 15a /2015. https://wiki.julkict.fi/julkict/avoin-data/dataportaali/avoin-tieto-ohry/kokous-15042015/avoimen-tiedon-vaikuttavuus-esitutkimus_suomi/view (Retrieved 4.11.2018)

Koski, H. Honkanen, M. Luukkonen, J. Pajarinen, M. Ropponen, T. 2017. Avoimen datan hyödyntäminen ja vaikuttavuus. Prime Minister's Office. Publications of the Government's analysis, assessment and research activities. 40/2017. <https://www.etla.fi/julkaisut/avoimen-datan-hyodyntaminen-ja-vaikuttavuus/> (Retrieved 4.11.2018)

Kreisberg, S. 1992. Transforming power: Domination, empowerment, and education. Albany, NY: State University of New York Press.

Lassinantti, J. Bergvall-Kåreborn, B. Ståhlbröst, A. 2013. Shaping Local Open Data Initiatives: Politics and Implications. Journal of Theoretical and Applied Electronic Commerce Research. Vol 9 / Issue 2.

Luttrell, C. Quiroz, S. Scrutton, C. Bird, K. 2009. Understanding and operationalizing empowerment. Working Paper 308. Overseas Development Institute. <http://demo.s.be/sites/default/files/5500.pdf> (Retrieved 4.11.2018)

Okamoto, K. 2016. What is being done with open government data? An exploratory analysis of public uses of New York City open data. Webology. 06/2016.

Page, N. Czuba, C. 1999. Empowerment: What is it? Journal of Extension 37 (5). <http://www.joe.org/joe/1999october/comm1.php> (Retrieved 4.11.2018)

Rappaport, J. 1984. Studies in Empowerment: Introduction to the Issue. Routledge, 2014.

Tuomi, J. Sarajärvi, A. 2003. Laadullinen tutkimus ja sisällönanalyysi. Tammi.

Van Dijk, J. 2012. The network society: Social aspects of new media. SAGE.

Zimmerman, M. 1990. Taking Aim on Empowerment Research: On the Distinction Between Individual and Psychological Conceptions. American Journal of Community Psychology. Vol. 18. No. 1.

<https://www.facebook.com/groups/fi.okfn/> (Retrieved 4.11.2018)

<https://www.europeandataportal.eu/en/providing-data/goldbook/open-data-nutshell> (Retrieved 4.11.2018)

<https://index.okfn.org/place/> (Retrieved 4.11.2018)

<https://okfn.org/opendata/> (Retrieved 4.11.2018)

<http://opendefinition.org/> (Retrieved 4.11.2018)

<http://opendefinition.org/government/> (Retrieved 4.11.2018)

<https://vm.fi/en/opendata> (Retrieved 4.11.2018)

Finnish open data portals that were contacted to promote the survey:

Avoindata.fi -portal: <https://www.avoindata.fi/fi> (Retrieved 4.11.2018)

HRI-portal: <http://www.hri.fi> (Retrieved 4.11.2018)

Lounaistieto-portal: <http://www.lounaistieto.fi/> (Retrieved 4.11.2018)

Oulu Open Data Portal: <https://data.ouka.fi/fi/> (Retrieved 4.11.2018)

Tampere Open Data Portal: <https://data.tampere.fi/fi/> (Retrieved 4.11.2018)

Turku Open Data: <https://www.turku.fi/avoindata> (Retrieved 4.11.2018)

Hei.

Teen parhaillaan pro gradu -tutkielmaa, jossa tarkastelen avoimen datan hyödyntämistä Suomessa erityisesti kansalaisen näkökulmasta. Tutkielmani puitteissa pyrin selvittämään, millä tavoin julkisen sektorin avaamaa dataa Suomessa hyödynnetään, sekä millaisia voimaannuttavia ja osallistavia kokemuksia datan käyttö synnyttää. Samalla olen kiinnostunut datan hyödyntämiseen liittyvistä motiiveista, sekä taidoista ja resursseista, joita datan käyttäjiltä vaaditaan.

Oheisen kyselyn avulla kerään tietoa avoimen datan merkityksistä, datan käyttötavoista, sekä tapausesimerkeistä, joissa avoimia tietoaaineistoja on hyödynnetty. Kyselyn tuloksiin pohjautuen tarkoitukseni on jatkaa aineistonkeruuta teemahaastatteluilla ja kyselyn loppupuolella voit ilmaista kiinnostuksesi myös haastattelua kohtaan.

Kyselyllä kerättyä tietoa käytetään ainoastaan pro gradu -tutkielmani tutkimustarpeisiin, tutkimuseettiset periaatteet huomioiden.

Kiitos vastauksestasi jo etukäteen!

Terveisin,

Aleksi Ruuhilahti
Viestintätieteiden tiedekunta
Tampereen yliopisto

Voit olla yhteydessä kyselyyn tai tutkimukseeni liittyvissä kysymyksissä s-postilla: ruuhilahti.aleksi.j@student.uta.fi

Kysely avoimen datan hyödyntäjille

Lomake on ajastettu: julkisuus alkaa 3.1.2018 14.15 ja päättyy 31.1.2018 23.59

Kiinnostuksesi avointa dataa kohtaan

1. Mitä seuraavista avointa dataa tarjoavista lähteistä olet käyttänyt?

Valtiolliset ja kunnalliset dataportaalit

- ☐ Avoindata.fi-portaali
- ☐ Helsinki Region Infoshare -portaali
- ☐ Lounaistieto-portaali
- ☐ Oulun kaupungin dataportaali
- ☐ Tampere Open Data -portaali
- ☐ Muu kunnallissektorin ylläpitämä datalähde

Muita julkisen sektorin ylläpitämiä avoimen datan lähteitä

- ☐ Evira
- ☐ Fingrid
- ☐ Geologian tutkimuskeskus
- ☐ Ilmatieteenlaitos
- ☐ Kansallisgalleria
- ☐ Liikennevirasto
- ☐ Maanmittauslaitos
- ☐ Patentti- ja rekisterihallitus
- ☐ Syke
- ☐ Terveystieteiden tutkimuskeskus
- ☐ Tilastokeskus
- ☐ Trafi
- ☐ Tukes
- ☐ Viestintävirasto

Muut käyttämäsi lähteet?

2. Arvioi seuraavia väittämiä sen perusteella, kuinka vahvasti ne kuvaavat käsitystäsi avoimesta datasta

	Ei ollenkaan	Heikosti	Osittain	Vahvasti
Avoin data on bittejä, joita on mahdollista hyödyntää ohjelmistokehityksessä ja uusissa innovaatioissa	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Avoin data on erityisesti talouden ja liike-toiminnan raaka-ainetta	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Avoimen datan avulla hallinnon toimintaa on mahdollista valvoa ja parantaa	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Avoimen datan käyttö parantaa kansalaisten yhteiskunnallisia vaikuttamismahdollisuuksia	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Hallinnon tuottaman tietoaaineiston vapaa saanti kuuluu perusoikeuksiin	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Voit kuvailla myös omin sanoin, mitä avoin data sinulle merkitsee.

3. Arvioi seuraavia väittämiä, jos pohdit mielenkiintoasi avointa dataa kohtaan tai rooliasi datan hyödyntäjänä.

	Ei ollenkaan	Heikosti	Osittain	Vahvasti
Avoim data liittyy työhöni	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Avoim data liittyy opintoihini	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Avoim data liittyy harrastuksiini tai vapaa-aikaani	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Avoim data liittyy aktiivisena kansalaisena toimimiseeni	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Vaikuttavatko jotkin muut tekijät kiinnostukseesi avointa dataa kohtaan?

Datan käyttö

4. Valitse seuraavista käyttöskenaarioista ne, jotka parhaiten kuvastavat tapojasi hyödyntää avointa dataa

- ☐ Olen tutkinut tietoaaineistoja hankkiakseni tietoa (esim. selailemalla tietoaaineistoa)
- ☐ Olen muodostanut tietoaaineistoihin perustuen uutta tietoa (esim. analysoimalla dataa ja luomalla analyysiin perustuvan visualisoinnin, raportin tai kirjoituksen)
- ☐ Olen luonut dataan pohjautuvan käyttöliittymän (esim. kehittämällä käyttöliittymän, jonka avulla voi etsiä tietoa tietoaaineistoon perustuen)
- ☐ Olen muokannut tai johtanut alkuperäisestä tietoaaineistosta uutta tietoaaineistoa (esim. korjaamalla tietoaaineistoa ja jakamalla sen eteenpäin toisessa formaatissa)
- ☐ Olen luonut dataan perustuvan palvelun tai hyödyntänyt dataa jo olemassa olevassa palvelussa

Jos kyseiset skenaariot eivät vastanneet kokemuksiasi tai olet käyttänyt dataa myös muilla tavoin, voit kuvailla käyttötapojasi alle.

5. Millaisiin aihepiireihin käyttämäsi tietoaaineistot ovat liittyneet? (esim. päätöksenteko, terveys, ympäristö yms.)

6. Mikä avoimen datan hyödyntämisessä sinua tällä hetkellä erityisesti kiinnostaa?

7. Millaista ongelmien/kysymysten ratkaisuun avoimen datan hyödyntäminen mielestäsi antaa mahdollisuuksia?

Tapausesimerkit

Tutkimukseni kannalta olisin kiinnostunut keräämään tietoa erilaisista tapausesimerkeistä, joissa avointa dataa on hyödynnetty. Esimerkit voivat liittyä niin isompiin projekteihin kuin yksittäisiin käyttötapauksiin. Jos mielessäsi on omakohtainen sellainen, voit kuvailla sitä alle.

Voit esimerkiksi pohtia seuraavia kysymyksiä: Mistä tapauksessa on/oli kyse ja miten avoimen datan hyödyntäminen siihen liittyi? Mitä tapauksen yhteydessä luotiin tai saavutettiin? Millaisia tietoaaineistoja tapauksen yhteydessä hyödynnettiin? Mitkä tekijät tapauksen taustalla vaikuttivat tai motivoivat?

Jos antamastasi esimerkistä löytyy lisätietoa internetistä, voit linkittää lähteitä oheen.

Osallistuminen haastatteluihin

Oletko kiinnostunut osallistumaan tutkimusaihepiiriä käsittelevään teemahaastatteluun? Kyllä En

Teemahaastatteluiden avulla pyrin keräämään lisäaineistoa tutkimukseni tueksi. Haastatteluiden painopiste tulee olemaan avointa dataa hyödyntävien henkilökohtaisissa kokemuksissa datan hyödyntämiseen liittyen, syventäen kyselyn aihepiiriä.

Pyrin järjestämään haastattelut alkukeväästä 2018.

Jos olet halukas osallistumaan haastatteluun, täytä alle mieluisin yhteydenottotapa (esim. s-postiosoite tai puh.)

Perustiedot

Sukupuoli Mies Nainen

Ikä alle 18 18-25 26-30 31-40 41-50 51-60 61-70 yli 70

Paikkakunta

Koulutus Peruskoulu Ammattitutkinto Lukio Alempi korkeakoulututkinto Ylempi korkeakoulututkinto Tohtori

Tieteenala

Mikä seuraavista kuvaa tämänhetkistä elämäntilannettasi parhaiten? Työssäkäyvä Työtön Opiskelija Eläkeläinen

Voit halutessasi kuvailla työskentelysektoriasi, työtehtävääsi tai osaamisaluetta tarkemmin alle.

Tietojen lähetyk

Tallenna

Järjestelmänä Eduix E-lomake 3.1, www.e-lomake.fi